



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/67, 15/30, 5/10, 1/21, C07K 14/445, A01K 67/027, A61K 48/00		A2	(11) International Publication Number: WO 99/20774 (43) International Publication Date: 29 April 1999 (29.04.99)		
(21) International Application Number: PCT/US98/22226		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).			
(22) International Filing Date: 20 October 1998 (20.10.98)					
(30) Priority Data: 60/062,592 20 October 1997 (20.10.97) US 60/085,649 15 May 1998 (15.05.98) US					
(71) Applicant: GENZYME TRANSGENICS CORPORATION [US/US]; 5 Mountain Road, Framingham, MA 01701 (US).					
(72) Inventors: CHEN, Li, How; 12 Wachusett Drive, Acton, MA 01720 (US). MEADE, Harry; 62 Grasmere Street, Newton, MA 01720 (US).					
(74) Agents: KEOWN, Wayne, A. et al.; Hale and Dorr LLP, 60 State Street, Boston, MA 02109 (US).					
(54) Title: NOVEL MODIFIED MSP-1 NUCLEIC ACID SEQUENCES AND METHODS FOR INCREASING mRNA LEVELS AND PROTEIN EXPRESSION IN CELL SYSTEMS					
(57) Abstract					
<p>The invention provides modified recombinant nucleic acid sequences (preferably DNA) and methods for increasing the mRNA levels and protein expression of malarial surface protein MSP-1 which is known to be difficult to express in cell culture systems, mammalian cell culture systems, or in transgenic animals. The preferred protein candidates for expression using the recombinant techniques of the invention are MSP-1 proteins expressed from DNA coding sequences comprising reduced overall AT content or AT rich regions and/or mRNA instability motifs and/or rare codons relative to the native MSP-1 gene.</p>					

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

NOVEL MODIFIED MSP-1 NUCLEIC ACID SEQUENCES
AND METHODS FOR INCREASING mRNA LEVELS AND PROTEIN
EXPRESSION IN CELL SYSTEMS

5

BACKGROUND OF THE INVENTION

10 Field of the invention

The invention relates to heterologous gene expression. More particularly, the invention relates to the expression of malaria genes in higher eukaryote cell systems.

15 Summary of the related art

Recombinant production of certain heterologous gene products is often difficult in *in vitro* cell culture systems or *in vivo* recombinant production systems. For example, many researchers have found it difficult to express proteins derived from bacteria, parasites and virus in cell culture systems different from the cell from 20 which the protein was originally derived, and particularly in mammalian cell culture systems. One example of a therapeutically important protein which has been difficult to produce by mammalian cells is the malaria merozoite surface protein (MSP-1).

25 Malaria is a serious health problem in tropical countries. Resistance to existing drugs is fast developing and a vaccine is urgently needed. Of the number of antigens that get expressed during the life cycle of *P. falciparum*, MSP-1 is the most extensively studied and promises to be the most successful candidate for vaccination. Individuals exposed to *P. falciparum* develop antibodies against 30 MSP-1, and studies have shown that there is a correlation between a naturally acquired immune response to MSP-1 and reduced malaria morbidity. In a number of studies, immunization with purified native MSP-1 or recombinant fragments of the protein has induced at least partial protection from the parasite (Diggs et al, (1993) *Parasitol. Today* 9:300-302). Thus MSP-1 is an important target for the

development of a vaccine against *P. falciparum*.

MSP-1 is a 190-220 kDa glycoprotein. The C-terminal region has been the focus of recombinant production for use as a vaccine. However, a major problem in 5 developing MSP-1 as a vaccine is the difficulty in obtaining recombinant proteins in bacterial or yeast expression systems that are equivalent in immunological potency to the affinity purified native protein (Chang et al., (1992) *J. Immunol.* **148**:548-555.) and in large enough quantities to make vaccine production feasible.

10 Improved procedures for enhancing expression of sufficient quantities of MSP-1 would be advantageous.

BRIEF SUMMARY OF THE INVENTION

The present invention provides improved recombinant DNA compositions and procedures for increasing the mRNA levels and protein expression of the malarial surface antigen MSP-1 in cell culture systems, mammalian cell culture systems, or in transgenic mammals. The preferred protein candidate for expression in an expression system in accordance with the invention is a C-terminal derivative of MSP-1 having a DNA coding sequence with reduced AT content, and eliminated mRNA instability motifs and rare codons relative to the recombinant expression systems. Thus, in a first aspect, the invention provides a DNA sequence derived from the sequence shown in SEQ ID NO 2. This derivative sequence is shown in SEQ ID NO 1.

In a second aspect, the invention provides a process for preparing a modified nucleic acid of the invention comprising the steps of lowering the overall AT content of the natural gene encoding MSP-1, eliminating all mRNA instability motifs and replacing all rare codons with a preferred codon of the mammary gland tissue, all by replacing specific codons in the natural gene with codons recognizable to, and preferably preferred by mammary gland tissue and which code for the same amino acids as the replaced codon. This aspect of the invention further includes modified nucleic acids prepared according to the process of the invention.

In a third aspect, the invention also provides vectors comprising modified MSP-1 nucleic acids of the invention and a goat beta casein promoter and signal sequence, and host cells transformed with nucleic acids of the invention.

In a fourth aspect, the invention provides transgenic non-human mammals whose germlines comprise a nucleic acid of the invention.

In a fifth aspect, the invention provides a DNA vaccine comprising a modified MSP-1 gene according to the invention.

DESCRIPTION OF THE DRAWINGS

Fig. 1 depicts the cDNA sequence of MSP-1₄₂ modified in accordance with the invention [SEQ ID NO 1] in which 306 nucleotide positions have been replaced to 5 lower AT content and eliminate mRNA instability motifs while maintaining the same protein amino acid sequence of MSP-1₄₂. The large letters indicate nucleotide substitutions.

Fig. 2 depicts the nucleotide sequence coding sequence of the "wild type" or native 10 MSP- 1₄₂ [SEQ ID NO 2].

Fig 3 is a codon usage table for wild type MSP-1₄₂ (designated "MSP wt" in the table) and the new modified MSP-1₄₂ gene (designated "edited MSP" in the table) and several milk protein genes (casein genes derived from goats and mouse). The 15 numbers in each column indicate the actual number of times a specific codon appears in each of the listed genes. The new MSP-1₄₂ synthetic gene was derived from the mammary specific codon usage by first choosing GC rich codons for a given amino acid combined with selecting the amino acids used most frequently in the milk proteins.

20

Fig. 4a-c depict MSP-1₄₂ constructs GTC 479, GTC 564, and GTC 627, respectively as are described in the examples.

Fig. 5 panel A is a Northern analysis wherein construct GTC627 comprises the new 25 MSP-1₄₂ gene modified in accordance with the invention, GTC479 is the construct comprising the native MSP-1₄₂ gene, and construct GTC469 is a negative control DNA

Fig 5 panel B is a Western analysis wherein the eluted fractions after affinity purifications numbers are collected fractions. The results show that fractions from

GTC679 the modified MSP-1₄₂ synthetic gene construct reacted with polyclonal antibodies to MSP-1 and the negative control GTC479 did not.

Fig 6 depicts the nucleic acid sequences of OT1 [SEQ ID NO 3], OT2 [SEQ ID NO 4],
5 MSP-8 [SEQ ID NO 5], MSP-2 [SEQ ID NO 6] and MSP1 [SEQ ID NO 7] described in the Examples.

Fig 7 is a schematic representation of plasmid BC574.

10 Fig 8 is a schematic representation of BC620.

Fig 9 is a schematic representation of BC670.

Fig 10 is a representation of a Western blot of MSP transgenic milk.

15

Fig 11 is a schematic representation of the nucleotide sequence of MSP42-2 [SEQ ID NO 8].

Fig 12 is a schematic representation of the BC-718.

20

Fig 13 is a representation of a Western blot of BC-718 expression in transgenic milk.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The patent and scientific literature referred to herein establishes the knowledge that is available to those with skill in the art. The issued US patents, 5 allowed applications, published foreign applications, and references cited herein are hereby incorporated by reference. Any conflicts between these references and the present disclosure shall be resolved in favor of the present disclosure.

The present invention provides improved recombinant DNA compositions 10 and procedures for increasing the mRNA levels and protein expression of the malarial surface antigen MSP-1 in cell culture systems, mammalian cell culture systems, or in transgenic mammals. The preferred protein candidate for expression in an expression system in accordance with the invention is a C-terminal derivative 15 of MSP-1 having a DNA coding sequence with reduced AT content, and eliminated mRNA instability motifs and rare codons relative to the recombinant expression systems. Thus, in a first aspect, the invention provides a DNA sequence derived from the sequence shown in SEQ ID NO 2. This derivative sequence is shown in SEQ ID NO 1.

In preferred embodiments, the nucleic acid of the invention is capable of 20 expressing MSP-1 in mammalian cell culture systems, or in transgenic mammals at a level which is at least 25%, and preferably 50% and even more preferably at least 100% or more of that expressed by the natural gene in mammalian cell culture systems, or in transgenic mammals under identical conditions.

As used herein, the term "expression" is meant mRNA transcription 25 resulting in protein expression. Expression may be measured by a number of techniques known in the art including using an antibody specific for the protein of interest. By "natural gene" or "native gene" is meant the gene sequence, or fragments thereof (including naturally occurring allelic variations), which encode the wild type form of MSP-1 and from which the modified nucleic acid is derived. 30 A "preferred codon" means a codon which is used more prevalently by the cell or

tissue type in which the modified MSP-1 gene is to be expressed, for example, in mammary tissue. Not all codon changes described herein are changes to a preferred codon, so long as the codon replacement is a codon which is at least recognized by the mouse mammary tissue. The term "reduced AT content" as used herein means 5 having a lower overall percentage of nucleotides having A (adenine) or T (thymine) bases relative to the natural MSP-1 gene due to replacement of the A or T containing nucleotide positions or A and/or T containing codons with nucleotides or codons recognized by mouse mammary tissue and which do not change the amino acid sequence of the target protein.

10

In a second aspect, the invention provides a process for preparing a modified nucleic acid of the invention comprising the steps of lowering the overall AT content of the natural gene encoding MSP-1, eliminating all mRNA instability motifs and replacing all rare codons with a preferred codon of mammary gland 15 tissue, all by replacing specific codons in the natural gene with codons recognizable to, and preferably preferred by mammary gland tissue and which code for the same amino acids as the replaced codon. Standard reference works describing the general principals of recombinant DNA technology include Watson, J.D. et al, *Molecular Biology of the Gene*, Volumes I and II the Benjamin/Cummings Publishing 20 Company, Inc. publisher, Menlo Park, CA (1987) Darnell, J.E. et al., *Molecular Cell Biology*, Scientific American Books, Inc., Publisher, New York, NY (1986); Old, R.W., et al., *Principles of Gene Manipulation: An Introduction to Genetic Engineering*, 2d edition, University of California Press, publisher, Berkeley CA (1981); Maniatis, T., et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor 25 Laboratory, publisher, Cold Spring Harbor, NY (1989) and *Current Protocols in Molecular Biology*, Ausubel et al., Wiley Press, New York, NY (1992). This aspect of the invention further includes modified nucleic acids prepared according to the process of the invention.

Without being limited to any theory, previous research has indicated that a

conserved AU sequence (AUUUA) from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation (Shaw, G. and Kamen, R. *Cell* 46:659-667). The focus in the past has been on the presence of these instability motifs in the untranslated region of a gene. The instant invention is the first to recognize 5 an advantage to eliminating the instability sequences in the coding region of the MSP-1 gene.

In a third aspect, the invention also provides vectors comprising modified 10 MSP-1 nucleic acids of the invention and a goat beta casein promoter and signal sequence, and host cells transformed with nucleic acids of the invention.

In a fourth aspect, the invention provides transgenic non-human mammals whose germlines comprise a nucleic acid of the invention. General principals for producing transgenic animals are known in the art. See for example Hogan et al., 15 *Manipulating the Mouse Embryo: A Laboratory Manual*, Cold Spring Harbor Laboratory, (1986); Simons et al, *Bio/Technology* 6:179-183, (1988); Wall et al., *Biol. Reprod.* 32:645-651, (1985); Buhler et al., *Bio/Technology*, 8:140-143 (1990); Ebert et al., *Bio/Technology* 9:835-838 (1991); Krimenfort et al., *Bio/Technology* 9:844-847 (1991); Wall et al., *J.Cell. Biochem.* 49:113-120 (1992). Techniques for introducing foreign 20 DNA sequences into mammals and their germ cells were originally developed in the mouse. See e.g., Gordon et al., *Proc. Natl. Acad. Sci. USA* 77:7380-7384, (1980); Gordon and Ruddle, *Science* 214: 1244-1246 (1981); Palmiter and Brinster, *Cell* 41: 343-345, 1985; Brinster et al., *Proc Natl. Acad Sci., USA* 82:4438-4442 (1985) and Hogan et al. (*ibid.*). These techniques were subsequently adapted for use with larger 25 animals including cows and goats. Up until very recently, the most widely used procedure for the generation of transgenic mice or livestock, several hundred linear molecules of the DNA of interest in the form of a transgenic expression construct are injected into one of the pro- nuclei of a fertilized egg. Injection of DNA into the cytoplasm of a zygote is also widely used. Most recently cloning of an entire

transgenic cell line capable of injection into an unfertilized egg has been achieved (KHS Campbell et al., *Nature* 380 64-66, (1996)).

The mammary gland expression system has the advantages of high expression levels, low cost, correct processing and accessibility. Known proteins, 5 such as bovine and human alpha-lactalbumin have been produced in lactating transgenic animals by several researchers. (Wright et al, *Bio/Technology* 9:830-834 (1991); Villette et al, *Eur. J. Biochem.*, 186:43-48 (1989); Hochi et al., *Mol Reprod. And Devel.* 33:160-164 (1992); Soulier et al., *FEBS Letters* 297(1,2):13-18 (1992)) and the system has been shown to produce high levels of protein.

10

In a fifth aspect, the invention provides a DNA vaccine comprising a modified MSP-1 gene according to the invention. Such DNA vaccines may be delivered without encapsulation, or they may be delivered as part of a liposome, or as part of a viral genome. Generally, such vaccines are delivered in an amount 15 sufficient to allow expression of the modified MSP-1 gene and to elicit an antibody response in an animal, including a human, which receives the DNA vaccine. Subsequent deliveries, at least one week after the first delivery, may be used to enhance the antibody response. Preferred delivery routes include introduction via mucosal membranes, as well as parenteral administration.

20

Examples

Creation of novel modified MSP-1₄₂ gene

25 A novel modified nucleic acid encoding the C-terminal fragment of MSP-1 is provided. The novel, modified nucleic acid of the invention encoding a 42 kD C-terminal part of MSP-1 (MSP-1₄₂) capable of expression in mammalian cells of the invention is shown in Fig. 1. The natural MSP-1₄₂ gene (Fig 2) was not capable of being expressed in mammalian cell culture or in transgenic mice. Analysis of the 30 natural MSP-1₄₂ gene suggested several characteristics that distinguish it from

mammalian genes. First, it has a very high overall AT content of 76%. Second, the mRNA instability motif, AUUUA, occurred 10 times in this 1100 bp DNA segment (Fig 2). To address these differences a new MSP-1₄₂ gene was designed. Silent nucleotide substitution was introduced into the native MSP-1₄₂ gene at 306

5 positions to reduce the overall AT content to 49.7%. Each of the 10 AUUUA mRNA instability motifs in the natural gene were eliminated by changes in codon usage as well. To change the codon usage, a mammary tissue specific codon usage table, Fig. 3a, was created by using several mouse and goat mammary specific proteins. The table was used to guide the choice of codon usage for the modified MSP-1₄₂ gene as

10 described above. For example as shown in the Table in Fig. 3a, in the natural gene, 65% (25/38) of the Leu was encoded by TTA, a rare codon in the mammary gland. In the modified MSP-1₄₂ gene, 100% of the Leu was encoded by CTG, a preferred codon for Leu in the mammary gland.

15 An expression vector was created using the modified MSP-1₄₂ gene by fusing the first 26 amino acids of goat beta-casein to the N-terminal of the modified MSP-1₄₂ gene and a SalI-Xho I fragment which carries the fusion gene was subcloned into the XhoI site of the expression vector pCDNA3. A His6 tag was fused to the 3' end of the MSP-1₄₂ gene to allow the gene product to be affinity purified. This resulted in

20 plasmid GTC627 (Fig.4c).

To compare the natural MSP-1₄₂ gene construct to the modified MSP-1₄₂ nucleic acid of the invention, an expression vector was also created for the natural MSP-1₄₂ gene and the gene was added to mammalian cell culture and injected into mice to

25 form transgenic mice as follows:

Construction of the native MSP-1₄₂ Expression Vector

To secrete the truncated merozoite surface protein-1 (MSP-1) of *Plasmodium falciparum*, the wild type gene encoding the 42KD C-terminal part of MSP-1 (MSP-1₄₂) was fused to either the DNA sequence that encodes the first 15 or the first 26 amino acids of the goat beta-casein. This is achieved by first PCR amplify the 5 MSP-1 plasmid (received from Dr. David Kaslow, NIH) with primers MSP1 and MSP2 (Fig. 6), then cloned the PCR product into the TA vector (Invitrogen). The Bg1II-XhoI fragments of the PCR product was ligated with oligos OT1 and OT2 (Fig. 6) into the expression vector pCDNA3. This yielded plasmid GTC564 (Fig.4b), which encodes the 15 amino acid beta- casein signal peptide and the first 11 amino acids of 10 the mature goat beta-casein followed by the native MSP-1₄₂ gene. Oligos MSP-8 and MSP-2 (Fig. 6) were used to amplify MSP-1 plasmid by PCR, the product was then cloned into TA vector. The XhoI fragment was excised and cloned into the XhoI site of the expression vector pCDNA3 to yield plasmid GTC479 (Fig.4a), which encoded 15 amino acid goat beta-casein signal peptide fused to the wild-type MSP-1₄₂ 15 gene. A His6 tag was added to the 3' end of MSP-1₄₂ gene in GTC 564 and GTC 479.

Native MSP-1₄₂ Gene Is Not Expressed In COS-7 Cells

20 Expression of the native MSP gene in cultured COS-7 cells was assayed by transient transfection assays. GTC479 and GTC564 plasmids DNA were introduced into COS-7 cells by lipofectamine (Gibco-BRL) according to manufacturer's protocols. Total cellular RNA was isolated from the COS cells two days post-transfection. The newly synthesized proteins were metabolically labeled for 10 hours by adding ³⁵S methionine added to the culture media two days-post transfection.

25

To determine the MSP mRNA expression in the COS cells, a Northern blot was probed with a ³²P labeled DNA fragment from GTC479. No MSP RNA was detected in GTC479 or GTC564 transfected (data not shown). Prolonged exposure revealed residual levels of degraded MSP mRNA. The ³⁵S labeled culture supernatants and

the lysates were immunoprecipitated with a polyclonal antibody raised against MSP. Immunoprecipitation experiments showed that no expression from either the lysates or the supernatants of the GTC479 or GTC564 transfected cells (data not shown). These results showed that the native MSP-1 gene was not expressed in COS

5 cells.

Native MSP-1₄₂ Gene is Not Expressed in the Mammary Gland of Transgenic Mice

10

The SalI-XhoI fragment of GTC479, which encoded the 15 amino acids of goat beta-casein signal peptide, the first 11 amino acids of goat beta-casein, and the native MSP-1₄₂ gene, was cloned into the XhoI site of the beta-casein expressed in vector BC350. This yielded plasmid BC574 (Fig.7). A SalI-NotI fragment of BC574 was

15 injected into the mouse embryo to generate transgenic mice. Fifteen lines of transgenic mice were established. Milk from the female founder mice was collected and subjected to Western analysis with polyclonal antibodies against MSP. None of the seven mice analyzed were found to express MSP-1₄₂ protein in their milk. To further determine if the mRNA of MSP-1₄₂ was expressed in the mammary gland,

20 total RNA was extracted from day 11 lactating transgenic mice and analyzed by Northern blotting. No MSP-1₄₂ mRNA was detected by any of the BC 574 lines analyzed. Therefore, the MSP-1₄₂ transgene was not expressed in the mammary gland of transgenic mice. Taken together, these experiments suggest that native parasitic MSP-1₄₂ gene could not be expressed in mammalian cells, and the block is

25 as the level of mRNA abundance.

Expression of MSP in the Mammalian Cells

Transient transfection experiments were performed to evaluate the expression of

the modified MSP-1₄₂ gene of the invention in COS cells. GTC627 and GTC479 DNA were introduced into the COS-7 cells. Total RNA was isolated 48 hours post-transfection for Northern analysis. The immobilized RNA was probed with ³²P labeled SalI-XhoI fragment of GTC627. A dramatic difference was observed 5 between GTC479 and GTC627. While no MSP-1₄₂ mRNA was detected in the GTC479 transfected cells as shown previously, abundant MSP-1₄₂ mRNA was expressed by GTC627 (Fig. 5, Panel A). GTC 469 was used as a negative control and comprises the insert of GTC564 cloned into cloning vector PU19, a commercially available cloning vector. A metabolic labeling experiment with ³⁵S methionine 10 followed by immunoprecipitation with polyclonal antibody (provided by D. Kaslow NIAID, NIH) against MSP showed that MSP-1₄₂ protein was synthesized by the transfected COS cells (Fig.5, Panel B). Furthermore, MSP-1₄₂ was detected in the transfected COS supernatant, indicating the MSP-1₄₂ protein was also secreted. 15 Additionally, using Ni-NTA column, MSP-1₄₂ was affinity purified from the GTC627 transfected COS supernatant.

These results demonstrated that the modification of the parasitic MSP-1₄₂ gene lead to the expression of MSP mRNA in the COS cells. Consequently, the MSP-1₄₂ product was synthesized and secreted by mammalian cells.

20 Polyclonal antibodies used in this experiment may also be prepared by means well known in the art (*Antibodies: A Laboratory Manual*, Ed Harlow and David Lane, eds. Cold Spring Harbor Laboratory, publishers (1988)). Production of MSP serum antibodies is also described in Chang et al., *Infection and Immunity* (1996) 64:253-261 25 and Chang et al., (1992) *Proc Natl. Acad. Sci. USA* 86:6343-6347.

The results of this analysis indicate that the modified MSP-1₄₂ nucleic acid of the invention is expressed at a very high level compared to that of the natural protein

which was not expressed at all. These results represent the first experimental evidence that reducing the AT % in a gene leads to expression of the MSP gene in heterologous systems and also the first evidence that removal of AUUUA mRNA instability motifs from the MSP coding region leads to the expression of MSP 5 protein in COS cells. The results shown in Fig. 5, Panel A Northern (i.e. no RNA with native gene and reasonable levels with a modified DNA sequence in accordance with the invention), likely explains the increase in protein production.

10 The following examples describe the expression of MSP1-42 as a native non-fusion (and non-glycosylated) protein in the milk of transgenic mice.

Construction of MSP Transgene

To fuse MSP1-42 to the 15 amino acid β -casein signal peptide, a pair of oligos, MSP203 15 and MSP204 (MSP203: ggccgctcgacgccaccatgaaggcttcataattgcc tgctggctggccattgcagccgtcactccctccgtcat, MSP204: cgatgacggaggagtgacggctg caatggccagagccaccagacaggcaattatgaggacccatggctggcgicgagc), which encode the 15 amino acid - casein signal and the first 5 amino acid of the MSP1-42 ending at the Cla I site, was ligated with a 20 Cla I-Xho I fragment of BC620 (Fig. 8) which encodes the rest of the MSP1-42 gene, into the Xho I site of the expression vector pCDNA3. A Xho I fragment of this plasmid (GTC669) was then cloned into the Xho I site of milk specific expression vector BC350 to generate B670 (Fig.9)

Expression of MSP1-42 in the milk of transgenic mice

25 A Sal I-Not I fragment was prepared from plasmid BC670 and microinjected into the mouse embryo to generate transgenic mice. Transgenic mice was identified by extracting mouse DNA from tail biopsy followed by PCR analysis using oligos GTC17 and MSP101 (sequences of oligos: GTC17, GATTGACAAGTAATACGCTGTTCCCTC, Oligo MSP101, GGATTCAATAGATACGG). Milk from the female founder transgenic mice was collected at day 30 7 and day 9 of lactation, and subjected to western analysis to determine the expression level of MSP-1-42 using an polyclonal anti-MSP antibody and monoclonal anti MSP antibody 5.2 (Dr. David Kaslow, NIH). Results indicated that the level of MSP-1-42 expression in the milk of

transgenic mice was at 1-2 mg/ml (Fig. 10).

Construction of MSP1-42 glycosylation sites minus mutants

5 Our analysis of the milk produced MSP revealed that the transgenic MSP protein was N-glycosylated. To eliminate the N-glycosylation sites in the MSP1-42 gene, Asn. (N) at positions 181 and 262 were substituted with Gln.(Q). The substitutions were introduced by designing DNA oligos that anneal to the corresponding region of MSP1 and carry the AAC to CAG mutations. These oligos were then used as PCR primers to produce DNA fragments that
10 encode the N to Q substitutions.

To introduce N262-Q mutation, a pair of oligos, MSPGYLYCO-3 (CAGGGAATGCTGCAGATCAGC) AND MSP42-2 (AATTCTCGAGTTAGTG
15 GTGGTGGTGGTGATCGCAGAAAATACCATG, FIG. 11), were used to PCR amplify plasmid GTC627, which contains the synthetic MSP1-42 gene. The PCR product was cloned into pCR2.1 vector (Invitrogen). This generated plasmid GTC716.

20 To introduce N181-Q mutation, oligos MSPGLYCO-1 (CTCCTTGTTCAGG AACTTGTAGGG) and MSPGLCO-2 (GTCCTGCAGTACACATATGAG, Fig 4) were used to amplify plasmid GTC 627. The PCR product was cloned into pCR2.1. This generated plasmid GTC700.

25 The MSP double glycosylation mutant was constructed by the following three steps: first, a Xho I-Bsm I fragment of BC670 and the Bsm I-Xho I fragment of GTC716 is ligated into the Xho I site of vector pCR2.1. This resulted a plasmid that contain the MSP-1-42 gene with N262-Q mutation. EcoN I-Nde I fragment of this plasmid was then replaced by the EcoN I-Nde I fragment from plasmid GTC716 to introduce the second mutation, N181-Q. A Xho I fragment of this plasmid was finally cloned into BC350 to generate BC718 (Fig. 12).

30 *Transgenic expression of nonglycosylated MSP-1*

BC718 has the following characteristics: it carries the MSP1-42 gene under the control of the β -casein promoter so it can be expressed in the mammary gland of the transgenic animal during lactation. Further, it encodes a 15 amino acid β -casein leader sequence fused directly to MSP1-42.

so that the MSP1-42, without any additional amino acid at its N-terminal, can be secreted into the milk. Finally, because the N-Q substitutions, the MSP produced in the milk of the transgenic animal by this construct will not be N-glycosylated. Taken together, the transgenic MSP produced in the milk by BC718 is the same as the parasitic MSP.

5 A SalI/XhoI fragment was prepared from plasmid BC718 and microinjected into mouse embryos to generate transgenic mice. Transgenic animals were identified as described previously. Milk from female founders was collected and analyzed by Western blotting with antibody 5.2. The results, shown in Figure 13, indicate expression of nonglycosylated MSP1 at a concentration of 0.5 to 1 mg/ml.

10

What is claimed is:

1. A modified MSP-1 nucleic acid sequence, as shown in SEQ ID NO 1.
- 5 2. A process for preparing a modified MSP-1 nucleic acid comprising the steps of lowering the overall AT content of the natural gene encoding MSP-1, eliminating all mRNA instability motifs and replacing all rare codons with a preferred codon of the mouse mammary gland.
- 10 3. A modified MSP-1 nucleic acid prepared by the process according to claim 2.
4. A vector comprising a modified MSP-1 nucleic acid according to claim 1 and a goat beta casein promoter and signal sequence.
- 15 5. A host cell transformed with the vector according to claim 4.
6. A transgenic non-human mammal comprising a vector according to claim 4.
7. A transgenic non-human mammal comprising a modified MSP-1 nucleic acid according to claim 1.
- 20 8. A DNA vaccine comprising a modified nucleic acid according to claim 1.

25

C

1 ccCgtCactccCtccgtCa GATAACATCctGtcCAA GATCgAGAA GGA GtA ~;
 19 Al a Val Thr Pro Ser Val Ile Asp Asn Ile Leu Ser Lys Ile Gl u Asn Gl u Tyr G
 56 AGGT GCT GtA CCT GAA Goc CCT GcAGG GGT Cta CCT GAG CCT GAA GAA GAA G
 19 Iu Val Leu Tyr Leu Lys Pro Leu Al a Gl y Val Tyr Arg Ser Leu Lys Lys Gl n
 109 CtGGA GAA CAA CGT GAT GAC Ctt CAA CGT GAA CGT GAG GAT AT CCT GAA CAGC
 37 Leu Gl u Asn Asn Val Met Thr Phe Asn Val Asn Val Lys Asp Ile Leu Asn Ser
 163 CGGtt CAA CAA GGG GGA GAA Ctt CAA GAA CGt CCT GGA GAGC GAT CtGAT Ccc
 55 Arg Phe Asn Lys Arg Gl u Asn Phe Lys Asn Val Leu Gl u Ser Asp Leu Ile Pr
 216 Cta CAA GgAT CtGac CAGCAGC AA Cta CGt GGT CAA GGA GAT CC Cta CAA GtC CC
 72 P o Tyr Lys Asp Leu Thr Ser Ser Asn Tyr Val Val Lys Asp Pro Tyr Lys Phe L
 269 tGAA CAA GGA GAA GAGAGATAA Gtt Ctg GAGAGTT CAA Cta Cat CAA GAGAT AG
 90 Leu Asn Lys Gl u Lys Arg Asp Lys Phe Leu Ser Ser Tyr Asn Tyr Ile Lys Asp Se
 324 Cat TgATAC CGAT AT CAA Ctt Ccc CAA CGAT GT CCT GGGATA Cta CAA GAT CCT
 108 P r Ile Asp Thr Asp Ile Asn Phe Al a Asn Asp Val Leu Gl y Tyr Tyr Lys Ile Le
 378 GtccGA GAA GtA CAA GAGC GAT CtGgAT TCA AT CAA GAA GtA Cat CAA AC GA TAA
 126 P u Ser Gl u Lys Tyr Lys Ser Asp Leu Asp Ser Ile Lys Lys Tyr Ile Asn Asp Ly
 432 GcAGGG AGA GAA CGAGAA GtAC CtGcc CTT CCT GAA CAA Cat CGAG ACC CtGtA
 144 P s Gl n Gl y Gl u Asn Gl u Lys Tyr Leu Pro Phe Leu Asn Asn Ile Gl u Thr Leu Ty
 486 CAA GAC CGt CAA CGATAA GATTGAT CtGt CGt GAT Cca CCT GGA Gcc CAA GGT
 162 P r Lys Thr Val Asn Asp Lys Ile Asp Leu Phe Val Ile His Leu Gl u Al a Lys Va
 Ndel
 540 Cct GAA Cta CAC ATAT GAGAA GAGC AAC GT GGA GGT CAA GAT CAA GGA GGT GAA
 180 P I Leu Asn Tyr Thr Tyr Gl u Lys Ser Asn Val Gl u Val Lys Ile Lys Gl u Leu As
 594 TTAC CtGAA GAC Cat CCA GGA TAA CCT TGGC CGATT CAA GAA GAA CAA CAA CCT
 198 P n Tyr Leu Lys Thr Ile Gl n Asp Phe Lys Leu Al a Asp Phe Lys Lys Asn Asn Asn Ph
 648 CGT CGG GAT Cgc CGAT CtGAGC Ac CGATT CAA ACC CAA CAA CAA CCT GAC CAA
 216 P e Val Gl y Ile Al a Asp Leu Ser Thr Asp Tyr Asn His Asn Asn Leu Leu Thr Ly
 702 GTT CCT GAG Cac CGGT ATGGT Ctt CGAAA CCT Ggc CAA GAC GGT CCT GAG CAA
 234 P s Phe Leu Ser Thr Gl y Met Val Phe Gl u Asn Leu Al a Lys Thr Val Leu Ser As
 756 CCT Gct GgAT GG GAA C GtGCA GGG GAT GtGAC AT CAGC CAA GcACCA GtG Tgt
 252 P n Leu Leu Asp Gl y Asn Leu Gl n Gl y Met Leu Asn Ile Ser Gl n His Gl n Cys Va
 810 GAA GAA Gca GtG TCC CCA GAA CAG CCG GtG TTT CAG AC CCT GgAT GA GAGAGA
 270 P I Lys Lys Gl n Cys Pro Gl n Asn Ser Gl y Cys Phe Arg His Leu Asp Gl u Arg Gl
 864 GGA GIG TAA GIG T CtGCT GAA CtACAA GCA GGA AGGT GATAA GtG TGT GgAAA C
 288 P u Gl u Cys Lys Cys Leu Leu Asn Tyr Lys Gl n Gl u Gl y Asp Lys Cys Val Gl u Asn
 919 CC CAAT CCT ACT TGT AAG GA GAA CAAT GGT GGT GAT GtG CCA GtG TAC CG
 307 P Pro Asn Pro Thr Cys Asn Gl u Asn Asn Gl y Gl y Cys Asp Al a Asp Al a Lys Cys Thr G
 977 AGGA GgAT TCA GG GAG CAA CAC GtG CAC CtGt GA GtG TAC CAA GcC TGATT
 326 P I u Gl u Asp Ser Gl y Ser Asn Gl y Lys Lys Ile Thr Cys Gl u Cys Thr Lys Pro Asp S
 1034 CCTT ATCC ACT GtT CGAT GGT AT CCT TGT G TAGT
 345 P er Tyr Pro Leu Phe Asp Gl y Ile Phe Cys Ser

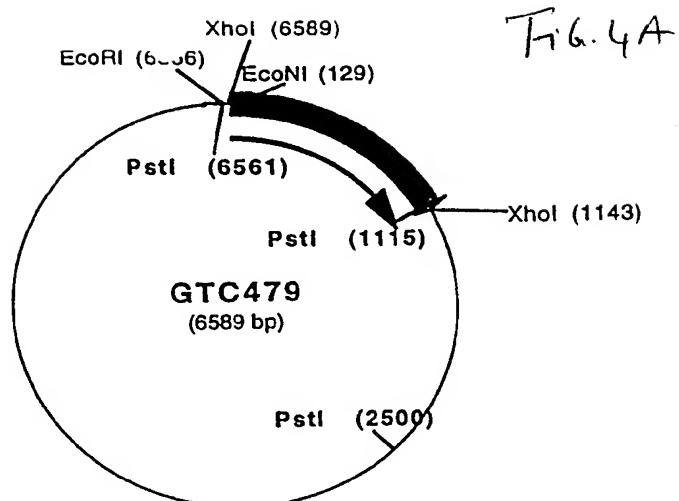
FIG. 1

1 GCAGTAACCTCCGTAAATT AACATACTTCTAAAGGATGAAATGAATP
 1► Al aVal Thr ProSer Val Ile Asn IleLeuSer Lys IleGluAsnGluTy
 EcoNI (73)
 56 AGGTTTTATATTTAAACCTTTAGCAGGTGTTATAGAAGTTAAAAAAACATT
 19► IuValLeuTyrLeuLysProLeuAlaGlyValTyrArgSerLeuLysLysGluLe
 111 AGAAAATAACGTTATGACATTAAATGTTAAGGATATTAAATTACCGA
 37► uGluAsnAsnValMetThrPheAsnValAsnValLysAspIleLeuAsnSerArg
 166 TTAAATAAACGTGAAATTCAAAATGTTAGAATCAGATTAAATTCCATATA
 56► PheAsnLysArgGluAsnPheLysAsnValLeuGluSerAspLeuIleProTyrL
 221 AAGATTTAACATCAAGTAATTATGTTGCAAGATCCATATAAATTCTTAATAA
 74► ysAspLeuThrSerSerAsnTyrValValLysAspProTyrLysPheLeuAsnLy
 276 AGAAAAAAAGAGATAAAATTCTTAAGCAGTTATAATTATAAAGGATCAATAGAT
 92► sGluLysArgAspLysPheLeuSerSerTyrAsnTyrIleLysAspSerIleAsp
 331 ACGGATATAAATTGCAAATGATGTTCTGGATAATTATAAAATTATCCGAAA
 111► ThrAspIleAsnPheAlaAsnAspValLeuGlyTyrTyrLysIleLeuSerGluLe
 386 AATATAAACAGATTCAATTAAAAAAATATAACGACAAACAAGGTGCA
 129► yStyRlysSerAspLeuAspSerIleLysStyRlysAsnAspLysGluGlyGlu
 441 AAATGAGAAATACTTCCCTTTAAACAATTGAGACCTTATAAAACAGTT
 147► uAsnGluLysTyrLeuProPheLeuAsnAsnIleGluThrLeuTyrLysThrVal
 496 AATGATAAAATTGATTATTGTAATTCAATTAGAAGCTCTAAATTATA
 166► AsnAspLysIleAspLeuPheValIleHisLeuGluAlaLysValLeuAsnTyrT
 551 CATATGAGAAATCAAACCTAGAAGTTAAAGAACTTAAATTACTTAAAC
 184► hrTyrGluLysSerAsnValGluValLysIleLysGluLeuAsnTyrLeuLysThr
 606 AATTCAAGACAAATTGGCAGATTAAACAAATTTCGTTGGAATTGCT
 202► rIleGlnAspLysLeuAlaAspPheLysLysAsnAsnAsnAsnAsnAsnAsnAsn
 661 GATTTATCAACAGATTATAACCATAATAACTTATTGACAAAGTTCCTTAGTACAG
 221► AspLeuSerThrAspTyrAsnHisAsnAsnLeuLeuThrLysPheLeuSerThrG
 716 GTATGGTTTTGAAAATCTTGCTAAAACGTTTATCTAATTACTTGATGGAAA
 239► yMeIValPheGluAsnLeuAlaLysThrValLeuSerAsnLeuLeuAspGlyAs
 771 CTGCAAGGTATGTTAACATTCAACACCAATGCGTAAAAAAACAAATGTCCA
 257► nLeuGlnGlyMeILeuAsnIleSerGlnHisGlnCysValLysLysGlnCysPro
 826 CAAAATTCTGGATGTTTCAGACATTAGATGAAAGAGAAGAATGTAATGTTAT
 276► GluAsnSerGlyCysPheArgHisLeuAspGluArgGluGluGlyCysLysCysLeuL
 881 TAAATTACAAACAAGAAGGTATAATGTTGAAATCCAAATCTACTTGTAA
 294► euAsnTyrLysGluGlyAspLysCysValGluAsnProAsnProThrCysAs
 936 CGAAAATAATGGTGGATGTCAGATGCCAAATGTCACCGAAGAAGATTCCAGGT
 312► nGluAsnAsnGlyGlyCysAspAlaAspAlaLysCysThrGluGluAspSerGly
 991 AGCAACGGAAAGAAAATCACATGTAATGTTACTAAACCTGATTCTTATCCACTTT
 331► SerAsnGlyLysLysIleThrCysGluCysThrLysProAspSerTyrProLeuP
 PstI (1059)
 1046 TCGATGGTATTTCTGAGTCACCAACCAACCACCAACTAACT
 349► heAspGlyIlePheCysSerHisHisHisHisHisHis

F16.2

Fig. 3A

Codon	AA	goat b-casein	goat K-casein	SP wt	Edited MSP	mouse b-casein	mouse a-casein	mouse	casein	mouse e-casein
TTT	Phe	5	4	8	0	4	8	3	3	4
TTC	Phe	4	3	7	15	4	6	7	7	1
TTA	Leu	0	2	25	0	0	0	0	0	0
TTG	Leu	0	2	3	0	0	0	0	0	1
TCT	Ser	5	1	4	1	13	5	7	7	5
TCC	Ser	2	2	2	3	6	14	8	8	2
TCA	Ser	1	4	10	1	1	3	2	2	0
TCG	Ser	0	1	0	0	0	0	0	0	0
TAT	Tyr	2	7	17	2	1	3	2	2	1
TAC	Tyr	1	2	3	18	2	6	6	6	7
TAA	---	1	2	0	0	1	0	1	0	0
TAG	---	0	0	0	0	0	0	0	0	0
TGT	Cys	1	1	10	12	0	0	1	1	0
TGC	Cys	0	2	2	0	2	2	2	2	1
TGA	---	0	0	0	0	0	1	0	0	1
TGG	Trp	1	1	0	0	0	2	2	2	2
CTT	Leu	9	1	9	0	16	9	3	3	3
CTC	Leu	5	2	0	0	7	6	0	0	1
CTA	Leu	1	2	1	0	1	2	1	1	0
CTG	Leu	11	5	0	38	10	17	4	4	1
CCT	Pro	17	6	4	2	8	6	3	3	0
CCC	Pro	12	0	1	6	8	6	6	6	4
CCA	Pro	3	13	5	1	5	6	2	2	2
CCG	Pro	1	1	0	1	0	0	0	0	1
CAT	His	0	1	3	0	2	6	2	2	1
CAC	His	5	3	1	4	4	0	3	0	0
CAA	Gln	5	9	9	0	9	21	9	9	7
CAG	Gln	16	6	0	9	21	32	12	12	8
CGT	Arg	0	1	1	0	0	0	0	0	0
CGC	Arg	0	0	0	0	1	0	0	0	0
CGA	Arg	0	0	1	0	0	0	0	0	1
CGG	Arg	1	0	0	3	0	0	0	0	0
ATT	Ile	4	5	13	0	3	4	3	3	4
ATC	Ile	6	3	2	20	7	5	8	8	5
ATA	Ile	1	3	5	0	1	0	2	2	0
ATG	Met	7	3	3	3	4	12	2	2	13
ACT	Thr	7	6	3	2	6	5	1	1	4
ACC	Thr	2	7	3	13	4	4	4	4	4
ACA	Thr	2	4	9	1	1	1	2	2	0
ACG	Thr	0	0	1	0	0	0	2	2	0
AAT	Asn	2	6	29	3	4	6	3	3	1
AAC	Asn	2	3	12	38	4	9	4	4	6
AAA	Lys	7	6	38	0	6	7	3	3	5
AAG	Lys	6	4	4	42	3	6	13	13	7
AGT	Ser	2	6	5	2	3	6	6	6	5
AGC	Ser	5	0	2	16	2	6	6	6	3
AGA	Arg	2	2	4	3	1	8	1	1	1
AGG	Arg	0	2	0	0	0	0	0	0	1
GTT	Val	5	6	15	0	7	4	2	2	3
GTC	Val	8	2	1	11	7	3	3	3	0
GTA	Val	2	2	5	0	2	4	1	1	3
GTG	Val	8	4	0	10	6	3	5	5	3
GCT	Ala	1	3	2	0	8	17	4	4	2
GCA	Ala	4	7	1	8	6	3	3	3	3
GCA	Ala	3	7	6	1	4	13	1	1	1
GCG	Ala	0	1	0	0	0	0	0	0	0
GAT	Asp	4	5	25	27	3	6	4	4	2
GAC	Asp	0	2	2	0	1	2	1	1	3
GAA	Glu	10	6	21	3	6	12	9	9	6
GAG	Glu	9	5	4	22	5	5	5	5	5
GGT	Gly	2	1	8	4	0	0	0	0	0
GGC	Gly	0	0	0	0	0	0	0	0	0
GGA	Gly	2	1	6	3	1	0	1	0	0
GGG	Gly	1	0	0	7	1	0	0	0	0



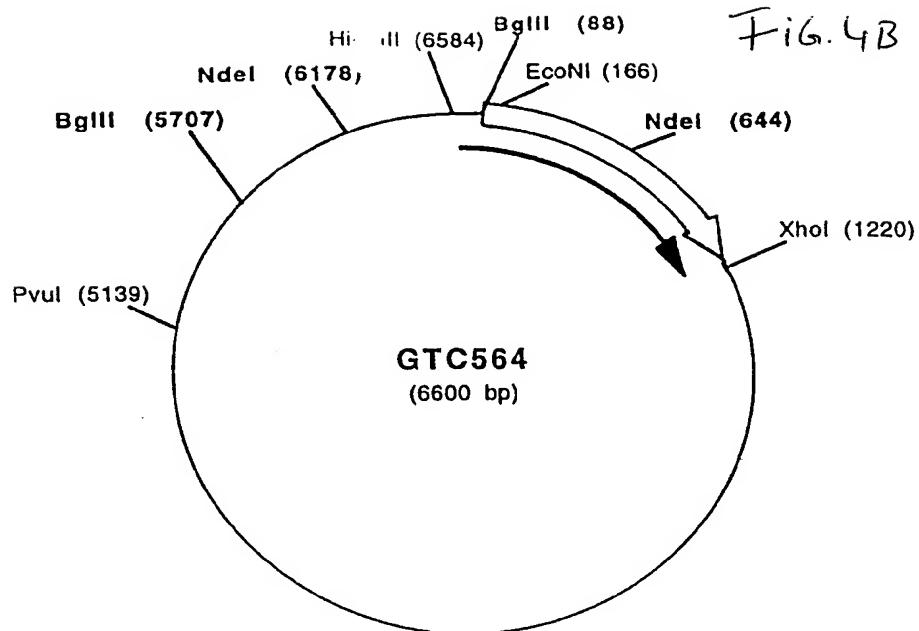


FIG. 4 C

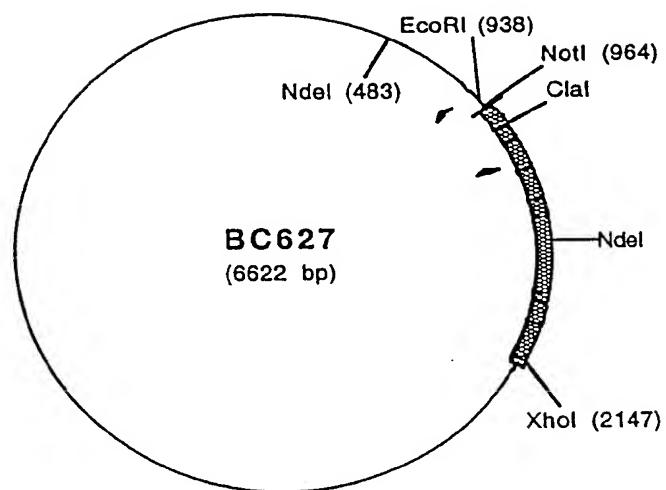
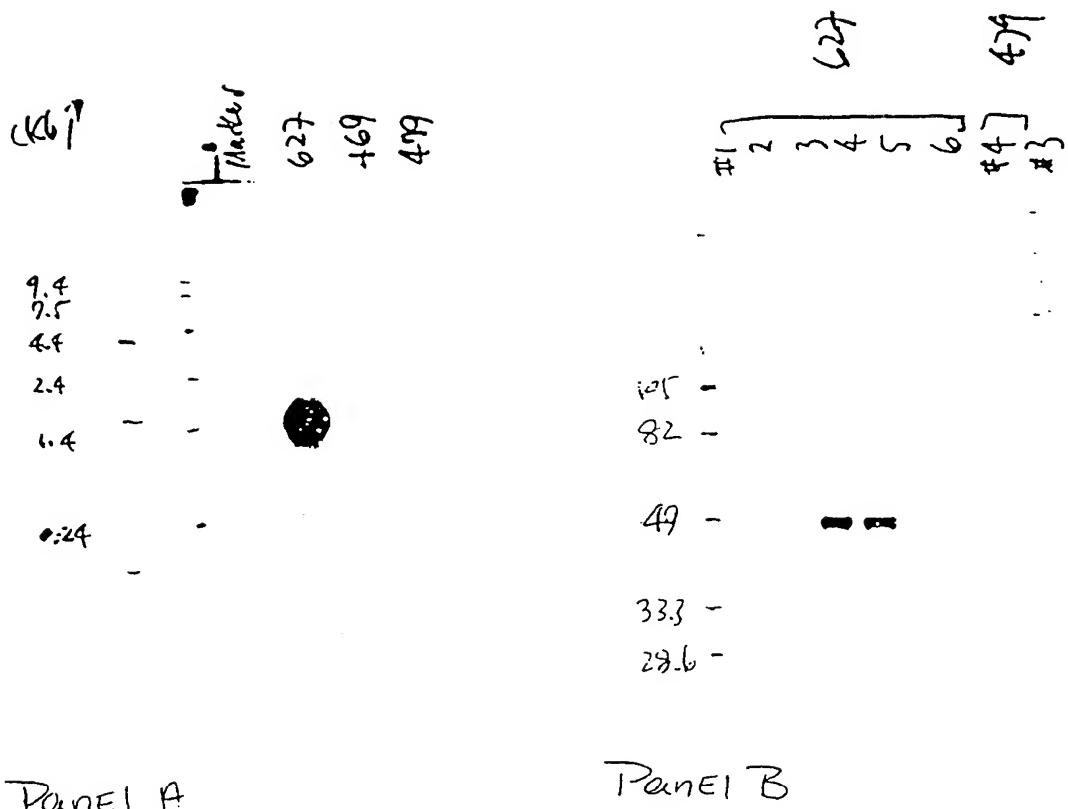


FIG. 5



Panel A

Panel B

Fig. 6

Oligos used:

OT1:

TCG ACG AGA GCC ATG AAG GTC CTC ATC CTT GCC TGT CTG GTG GCT
CTG GCC ATT GCA AGA GAG CAG GAA CTC AAT GTA GTC GGT A,

OT2:

GAT CTA CCG ACT ACA TTG AGT TCT TCC TGC TCT CTT GCA ATG GCC
AGA GCC ACC AGA CAG GCA AGG ATG AGG ACC TTC ATG GCT CTC G,

MSP1:

AATAGATCTGCAGTAACTCCTCCGTAATTG,

MSP2:

AATTCTCGAGTTAGTGGTGGTGGTGGTGGTACTGCAGAAATACCATC

MSP8:

TAACTCGAGCGAACCATGAAGGTCCATCCTGCCTGTCTGGTGGCTCTGG
CCATTGCA

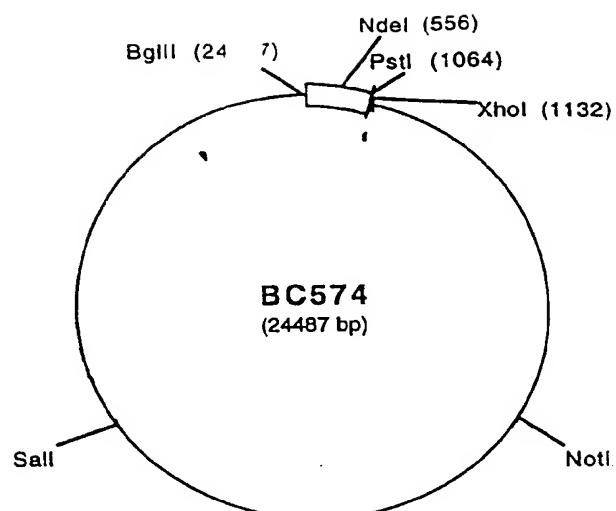


FIG. 7

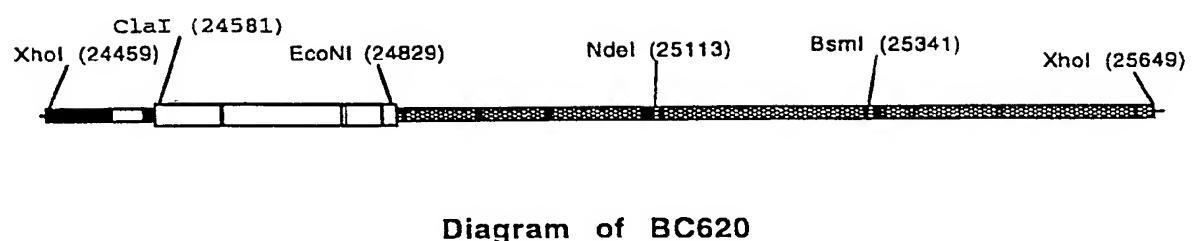


FIGURE 8

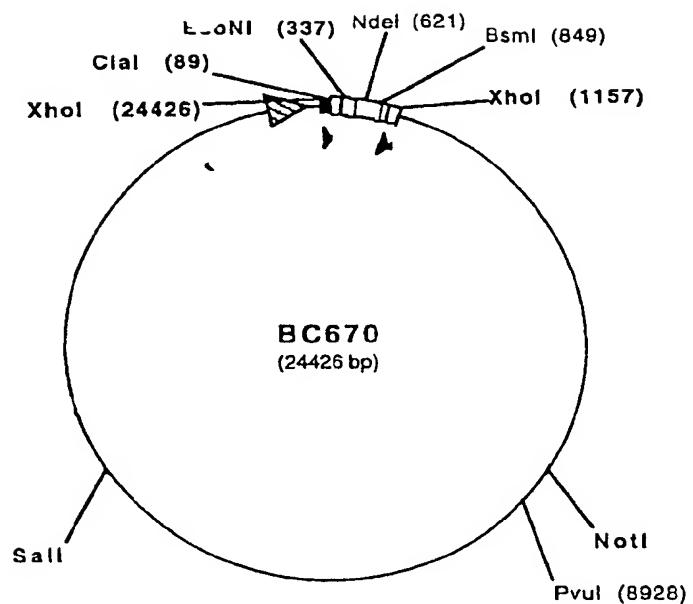
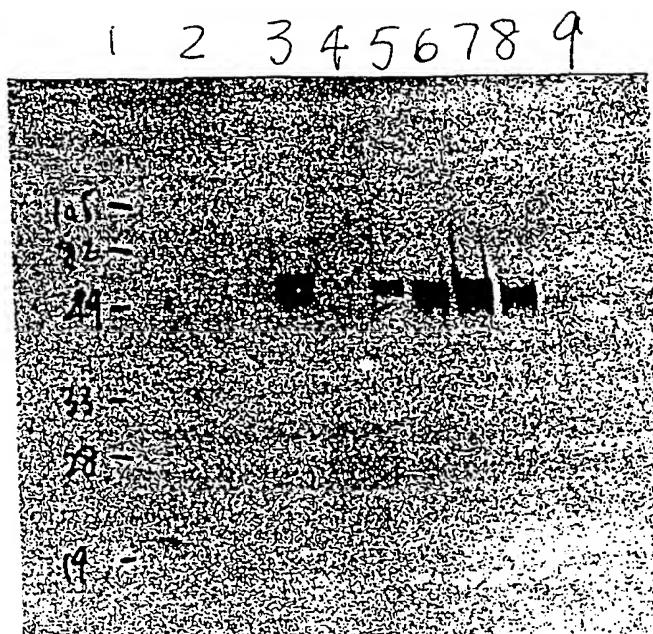


FIGURE 9

**Western Analysis of MSP transgenic milk.**

Lane 1, Molecular weight marker; lane 2, nontransgenic mice milk; lane 3, milk from BC628-146 transgenic mouse; lane 4-9, milk from BC670 transgenic mice. The blot was reacted with monoclonal antibody 5.2 against MSP.

FIGURE 10

26 ATGAAAGGTCTCTATAATTGCCCTGTC... GTGGCTCTGGGCAATTGGCAGCCGTCACTCCCTCCGTATCGATAAC
 1 M K V L I I A C L V A L A I A A V T P S V I D N
 98 ATCCCTGTCAGATCGAGAACCGAGTACCGAGGTCCTGACTCTGAAGGCCCTGGCAGGAGTCACAGGAGCCT
 25 I L S K I E N E Y E V L Y L K P L A G V Y R S L
 169 GAAGAACCGAGCTGGAGAACAACTGATGACCTTCACCTGAACCTGAAGGATATCTGAAACAGCAGGTCAAA
 46 K K O L E N N V M T F N V N V K D I L N S R F N
 241 CAAGAGGGAGAACCTCAAGAACGTGCTGGAGAGCGATCTGATCCCTACAGGATCTGACCAAGCAGGTC
 72 K R E N F K N V L E S D L I P Y K D L T S S N Y
 313 CGTGGTCAAAAGATCCCTACAACTTCTGAACAAGGAGAACAGAGATAAGTCCCTGACCAAGTACAAATTACAT
 96 V V K D P Y K F L N K E K R D K F L S S Y N Y I
 385 CAAGGATAGCATGGACACCGATATCAACCTTGCGAACGATGTCCTGGATACTACAAAGATCCCTGTCGAGAA
 120 K D S I D T D I N F A N D V L G Y Y K I L S E K
 457 GTACAAAGAGCCATCTGGATAGCATCAAGAAATACATCAACGATAAGCAGGGAGAGAACGGAAATCTGCG
 144 Y K S D L O S I K K Y I N D K Q G E N E K Y L P
 529 CCTCTCTGAAACACATCGAGACCCCTGACAAAGACCGTCACCGATAAGATTCATCTGATCTGATCCACCTGG
 168 F L N N I E T L Y K T V N D K I D L F V I H L E
 601 GGGCAAGGTCTTCGCAAGTACACATATGAGAGAGGAAACGTTGGAGGTCAGATCAAGGAGCTGAATTACCTGAA
 192 A K V L Q Y T Y E K S N V E V K I K E L N Y L K
 673 GACCATCCAGGATAAGCTGGCGGRTTCTCAAGAAGAACAAACTCTGGTCGAAATCGCGATCTGAGCACCGA
 216 T I Q D K L A D F K K N N N F V G I A D L S T D
 745 TTACAACCCACAACAAACCTGCTGACCAAGTCTCTGAGCACCGGAAATGGCTCTGGAAAACCTGGCCAAGACCGT
 240 Y N H N N L L T K F L S T G M V F E N L A K T V
 817 CCTGAGCAACCTGCTGGATGQAACCTGCAAGGAAATGCTGCAAGTCAGCAGCACCGAGTGTGAGAACAGC
 264 L S N L L D G N L Q G M L O I S Q H Q C V K K
 888 AGTGTCCCCAGAACACAGCGGATGCTTCAGAACACCTGGATGAGAGGGAGGTGCAAGTGCCTGCTGAACTA
 288 Q C P Q N S G G C F R H L D E R E E C K C L L N Y
 958 CAACAGGAGGAGGATAAGTGCTGGAACACCCCAATCTACTTGTAACGAGAACATGGAACATGGAACATGCGATG
 311 K Q E G D K C V E N P N P T C N E N N G G C D
 1029 CGATGCCAAGTGTACCGAGGAGGATTCAAGAACCAACGGAAAGAAGATCACCTGCCAGTGTACCAAGCCT
 335 A D A K C T E E D S G S N G K K I T C E C T K P
 1100 GATTCTTATCCACTGTCGATGATTCCTGCACTGACCCACCCACCAACTAACCTGCGAGGAT
 359 D S Y P L F D G I F G C S H H H H H H L E D

FIGURE 11

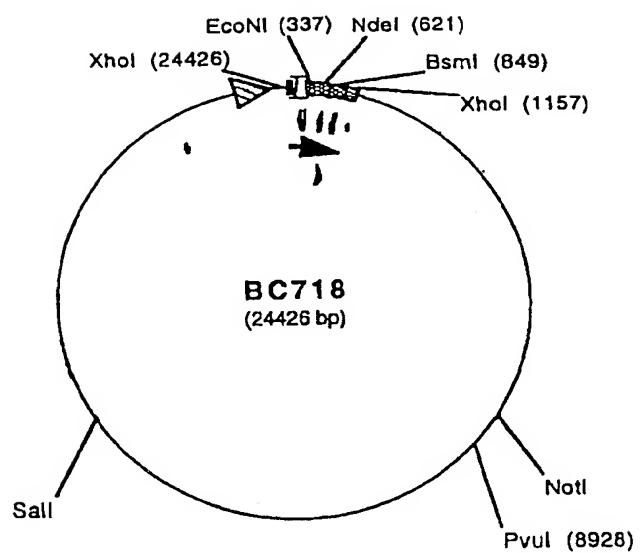


FIGURE 12

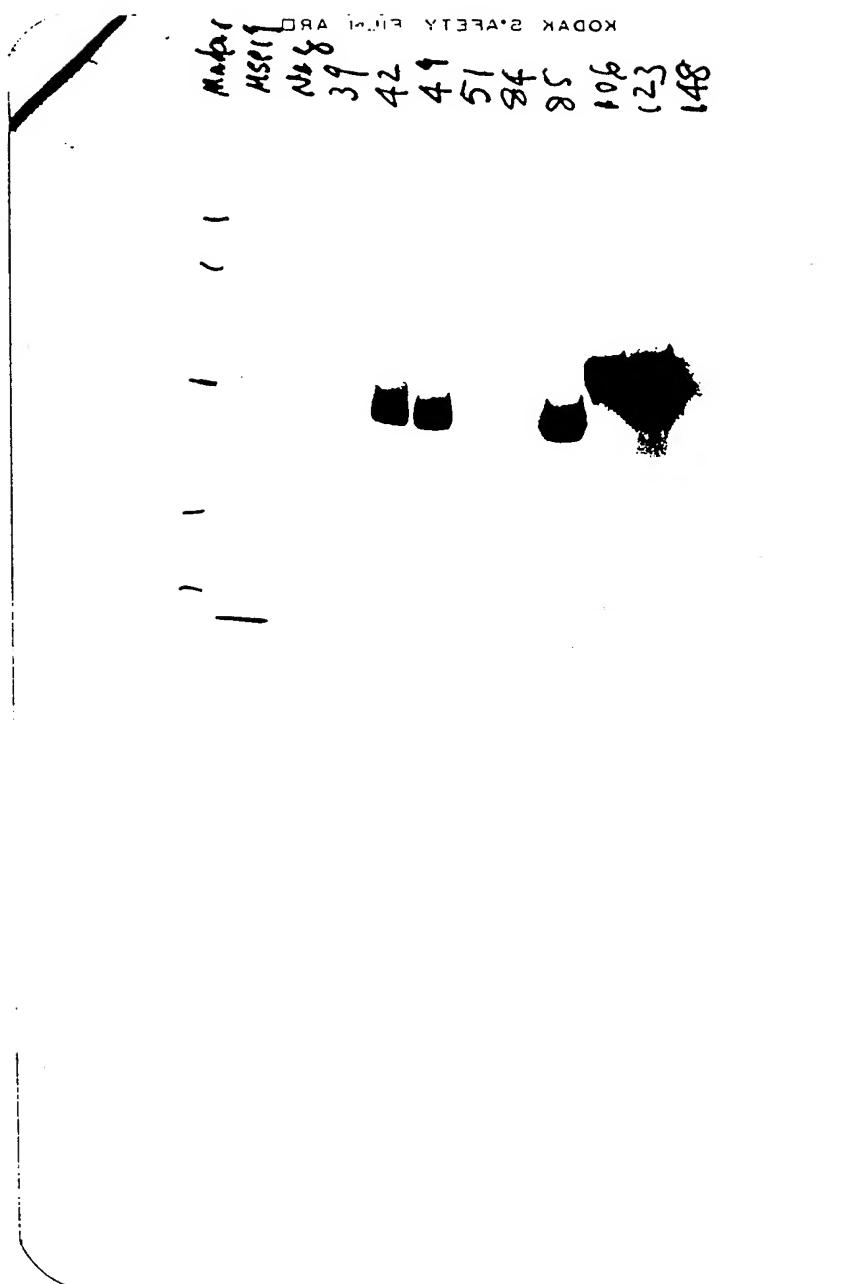


FIGURE 13